

Differentially Variable Component Analysis (dVCA): Identifying Multiple Evoked Components using Trial-to-Trial Variability

Kevin H. Knuth¹, Ankoor S. Shah^{2,3}, Wilson Truccolo⁴, Mingzhou Ding⁵,
Steven L. Bressler⁵, Charles E. Schroeder^{2,3}

1 Computational Sciences Division, Code IC, NASA Ames Research Center, Moffett
Field CA 94035

2 Department of Neuroscience, Albert Einstein College of Medicine, Bronx NY 10461

3 Cognitive Neuroscience and Schizophrenia Program, The Nathan S. Kline Institute,
Orangeburg, NY 10962

4 Neuroscience Department, Brown University, Providence RI 02912

5 Center for Complex Systems and Brain Sciences, Florida Atlantic University, Boca
Raton FL 33431

Phone: +1-650-604-4279

Fax: +1-650-604-4036

Email: kevin.h.knuth@nasa.gov

Abbreviated title: Differentially Variable Component Analysis (dVCA)

Number of Figures: 8

Number of tables: 2

Key words: Source separation, ICA, EEG, MEG, ERP, evoked responses, Bayesian
methods

October 8, 2003

SUMMARY

Electric potentials and magnetic fields generated by ensembles of synchronously active neurons in response to external stimuli provide information essential to understanding the processes underlying cognitive and sensorimotor activity. Interpreting recordings of these potentials and fields is difficult as each detector records signals simultaneously generated by various regions throughout the brain. We introduce the differentially Variable Component Analysis (dVCA) algorithm, which relies on trial-to-trial variability in response amplitude and latency to identify multiple components. Using simulations we evaluate the importance of response variability to component identification, the robustness of dVCA to noise, and its ability to characterize single-trial data. Finally, we evaluate the technique using visually evoked field potentials recorded at incremental depths across the layers of cortical area V1, in an awake, behaving macaque monkey.

INTRODUCTION

The field of neuroelectrophysiology relies on the analysis of electric potentials or magnetic fields produced by the brain in response to sensory stimulation, or in association with its cognitive and/or motor operations. These signals arise from transmembrane current flow produced by multiple ensembles of synchronously firing neurons. Far from being independent, these neural ensembles, also referred to as generators or sources, are often dynamically coupled in unknown ways that are of interest to the experimenter. Unfortunately, recording channels such as electrodes in electroencephalography (EEG) and superconducting quantum interference devices (SQUIDS) in magnetoencephalography (MEG) record linear mixtures of the activity from these sources in addition to ongoing background activity and sensor noise. Thus, the individual responses of each source are mixed within the recorded signal making it difficult to identify them and study their dynamical interactions. Furthermore, it is standard practice to enhance the signal-to-noise ratio by averaging event-related potentials (ERPs) over a number of experimental trials. However, implicit in this construction is the assumption that the evoked waveform is constant over trials and that any variability represents noise. In this practice, the possibility of assessing trial-dependent effects in the data is sacrificed.

The last decade has seen great developments in linear blind source separation (BSS) and independent component analysis (ICA) techniques, such as Infomax ICA (Bell & Sejnowski, 1995), FastICA (Hyvärinen & Oja, 1997), and second-order blind identification (SOBI) (Belouchrani et al., 1993). These algorithms have been useful in identifying sources in EEG and MEG signals using both ensemble-averaged data (Makeig et al., 1997; Särelä et al., 1998; Vigário et al., 1999) and single trials (Jung et al., 1999; Cao et al., 2000; Makeig et al., 2002; Tang et al., 2002). However, with the exception of SOBI, the general assumption that the ERP sources are independent is physiologically implausible. It is hard to argue that activity produced by two areas, such as V1 and V4, are independent when the areas are clearly structurally and functionally interconnected. Most important, by assuming independence of the sources, the experimenter assumes away one of the most interesting and important aspects of the behavior of active neural ensembles in the brain: the nature of their dynamical interactions.

In this paper we describe a more realistic model of the evoked response, which explicitly acknowledges trial-to-trial amplitude and latency variability in the evoked responses generated by each active source. Using this more realistic model of the ERP we will derive the differentially Variable Component Analysis (dVCA) algorithm, and demonstrate how different variability patterns in neural ensemble activity can be used to separate and identify their component signals. Using simulations, we will evaluate the ability of dVCA to characterize single-trial data, as well as its robustness to noise. Last, we will demonstrate how to use dVCA by analyzing visual-evoked field potentials recorded intracortically in macaque V1 using a linear multi-electrode array. Throughout this paper we will demonstrate that this approach allows us to: (1) more accurately account for observed trial-to-trial variability, (2) utilize the differential variation in the amplitudes and latencies to separate and identify sources, (3) avoid enforcing statistical independence of the sources, and (4) more accurately estimate the ongoing background activity.

MODELING EVOKED RESPONSES

Trial-to-trial variability of evoked responses can conceivably take many forms: latency shifts, amplitude variation, and even waveshape changes. In our experience, much of the observed variability can be described by amplitude and latency variations of a stereotypic response waveshape (Truccolo et al., 2002). For this reason, we model the response evoked from a single neural ensemble by assuming that the signal possesses a stereotypic waveshape that can vary in amplitude and onset latency from trial to trial. We write the response evoked in a given trial mathematically as $\alpha s(t - \tau)$, where the function $s(\cdot)$ represents the waveshape of the response as a function of elapsed time t , α represents the trial-specific amplitude scaling factor of the response, and τ represents the trial-specific onset latency shift. Furthermore, when multiple neural ensembles are engaged in the response to a stimulus, the activity of each ensemble is represented in the model as a separate waveform with a distinct trial-specific amplitude and latency shift. It is important to note that by modeling both the waveshape and its amplitude and latency, there is degeneracy in the model as an overall change in amplitude scale or latency shift can be described either by the amplitude and latency parameters or by the overall scale and shift of the waveshape. To eliminate this degeneracy, we take as a convention that the ensemble average amplitude scaling factor of the response over the recorded trials is unity, $\langle \alpha \rangle = 1$, and the ensemble average latency shift is zero, $\langle \tau \rangle = 0$.

In many experimental paradigms, investigators use several detectors positioned at different locations to measure the evoked responses of neural ensembles. The degree to which a detector records a signal evoked by a particular source depends on many factors including the relative position and orientation of the source relative to the detector. To describe this source-detector coupling, we introduce a coupling matrix \mathbf{C} , where the matrix element C_{mn} describes the degree to which the m^{th} detector detects the n^{th} source. This coupling matrix is known as the mixing matrix in the source separation literature and as the lead-field matrix in electrophysiology.

During the course of an experiment the investigator records responses to multiple presentations of a stimulus. Each presentation is typically called a trial. For the r^{th} recorded trial, we model the data $x(t)$ recorded by the m^{th} detector in component form as

$$\hat{x}_{mr}(t) = \sum_{n=1}^N C_{mn} \alpha_{nr} s_n(t - \tau_{nr}) + \eta_{mr}(t), \quad (1)$$

where $\hat{x}_{mr}(t)$ describes our model of the recorded data $x_{mr}(t)$, n indexes the N neural sources activated by stimulus presentation, C_{mn} is the coupling between the m^{th} detector and the n^{th} source, α_{nr} is the amplitude scale of the n^{th} source during the r^{th} trial, τ_{nr} is the latency shift of the n^{th} source during the r^{th} trial, $s_n(\cdot)$ is the waveshape of the n^{th} source, and $\eta_{mr}(t)$ is the un-modeled part of the data recorded in the m^{th} detector during the r^{th} trial. The un-modeled part of the data record is typically a combination of the recorded background activity along with any noise in the detector. For simplicity, we assume that $\eta_{mr}(t)$ has zero mean. Thus (1) describes the data recorded in a given detector during a given trial as a sum of the signals generated by each of the N neural

sources appropriately scaled in amplitude and shifted in latency for that trial and also scaled according to the coupling between each source and that detector plus the signals that we do not yet understand or care to understand. We call (1) the *multiple component event-related potential* (mcERP) model of evoked activity.

By adopting the mcERP model of the evoked responses, we implicitly adopt a well-defined set of characteristics capable of describing a neural source. The term component refers to the stereotypic waveshape describing the temporal activation pattern of a particular neural source. As no information regarding the spatial locations or distributions of the neural sources has gone into the model, this model does not distinguish between two spatially distinct groups of neurons that produce the same stereotypic waveshape varying identically with latency and proportionally in amplitude. However, in such a situation, examination of the estimated coupling matrix would reveal two spatially distinct sources if there are detectors positioned within the proximity of each source. The major advantage obtained by estimating the coupling matrix is that practical experience in conjunction with previously obtained physiological or anatomical data suggesting source distributions can be utilized to independently evaluate the solutions obtained with this technique. The disadvantage, aside from withholding information from the algorithm, is that there remain two degeneracies in the model. First, there is no specified order to the sources. Second, there is no specified scale for the coupling matrix; one could halve the coupling matrix elements while doubling the magnitude of the source waveshapes and obtain an equally valid solution. These degeneracies, which are present in every other linear source separation algorithm, pose no difficulties to the interpretation of the results. In our implementation we have chosen to normalize the columns of the coupling matrix so that the maximum value is equal to one. However, it should be noted that this scaling degeneracy could be remedied by adopting a physical model of the source currents (Knuth & Vaughan, 1998).

One last notable benefit of the mcERP model is that no component waveform is required to be present in every trial. In other words, a single-trial amplitude of zero is allowable for any component. By not assuming *a priori* that identical neural processes are active during every trial of the experiment, one is able to investigate the possibility that different processing strategies are used during the course of the recordings.

ALGORITHM DEVELOPMENT

Bayes' Theorem is the natural starting point for explaining the dVCA algorithm because it allows one to describe the probability of the model in terms of the likelihood of the data and the prior probability of the model parameters

$$p(model | data, I) = \frac{p(data | model, I) p(model | I)}{p(data | I)}, \quad (2)$$

where I represents any prior information one may have about the physical situation. The probability on the left-hand side of (2) is referred to as the posterior probability. It represents the probability that a given set of hypothesized values of the model parameters accurately describes the physical situation in which the data were collected. The first term in the numerator on the right-hand side is the likelihood of the data given the model. It describes the degree of accuracy with which we believe the model can predict the data.

The second term in the numerator is the prior probability of the model, or the prior. This prior describes the degree to which we believe the hypothesized values of the model parameters to be correct based only on our prior information about the problem. The term in the denominator is called the evidence and as we will see in this problem, it acts only as a normalization factor. It is through the assignment of the likelihood and the priors that we express all of our knowledge about the particular source separation problem. Bayes' Theorem can be viewed as describing how one's prior probability, $p(\text{model} | I)$, is modified by the acquisition of some new information in the form of data.

To apply this theorem to our problem, we consider the change in our knowledge about the model with the acquisition of new data, which consists of the set of recorded data $\mathbf{x}(t)$ from M detectors over the course of R trials. In this case, Bayes' Theorem can be written as

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) = \frac{p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | I)}{p(\mathbf{x}(t) | I)}, \quad (3)$$

where boldface symbols represent the entire set of parameters of each type in the mcERP model, eg. $\mathbf{a} = \{\alpha_1, \alpha_2, \dots, \alpha_R\}$. The most probable set of model parameters maximizes the probability in (3), and thus in practice the equation can be expressed as a proportionality with the inverse of the evidence $p(\mathbf{x}(t) | I)$ as the implicit proportionality constant. Equation (3) then becomes

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | I). \quad (4)$$

For simplicity, the joint prior can be factored into four terms each describing what we know about the source-detector coupling, the source waveshapes, the single-trial amplitudes and the single-trial latency offsets,

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto p(\mathbf{x}(t) | \mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, I) p(\mathbf{C} | I) p(\mathbf{s}(t) | I) p(\mathbf{a} | I) p(\boldsymbol{\tau} | I). \quad (5)$$

For the amplitude and latency priors, $p(\mathbf{a} | I)$ and $p(\boldsymbol{\tau} | I)$ respectively, we assign uniform densities with appropriate cutoffs denoting a range of physiologically realizable values. Note that a joint uniform density $p(\mathbf{a}, \boldsymbol{\tau} | I)$ can always be factored in this way. As the amplitude and latency priors are each represented by a uniform density, we can absorb those two factors into the implicit proportionality constant.

Our derivation continues by utilizing the principle of maximum entropy to assign a Gaussian likelihood (eg. Sivia, 1996; Jaynes, 2003) by introducing a parameter σ reflecting the expected square-deviation between our predictions and the mean

$$p(\mathbf{C}, \mathbf{s}(t), \mathbf{a}, \boldsymbol{\tau}, \sigma | \mathbf{x}(t), I) \propto (2\pi\sigma^2)^{-\frac{MRT}{2}} \text{Exp}\left[-\frac{1}{2\sigma^2} Q\right] p(\sigma | I) p(\mathbf{C} | I) p(\mathbf{s} | I), \quad (6)$$

where $p(\sigma | I)$ is the prior probability for σ . Q represents the sum of the square of the residuals between the data and our model in (1), written

$$Q = \sum_{m=1}^M \sum_{r=1}^R \sum_{t=1}^T \left(x_{mr}(t) - \sum_{n=1}^N C_{mn} \alpha_{nr} s_n(t - \tau_{nr}) \right)^2, \quad (7)$$

with M representing the number of detectors, R the number of experimental trials, and T the number of recorded time points per trial. Some consider this a fancy way to state that the noise is Gaussian distributed. However, as Jaynes demonstrates (2003), this is only a statement that the variance of the noise is equal to some value σ^2 . Thus higher-order statistical structure in the noise is still acceptable and the noise does not have to be Gaussian distributed.

The fact that we do not actually know the value of σ is not a problem as we can obtain a conservative result by considering all possible noise levels by integrating the joint posterior over all possible values of σ . Symmetry considerations require that we assign what is called a Jeffreys prior for σ , $p(\sigma | I) = \sigma^{-1}$ (Sivia, 1996). Performing the integral of the joint posterior over all possible values of σ , we obtain a marginal posterior probability for our original set of model parameters

$$p(\mathbf{C}, \mathbf{s}(t), \boldsymbol{\alpha}, \boldsymbol{\tau} | \mathbf{x}(t), I) \propto Q^{-\frac{MRT}{2}} p(\mathbf{C} | I) p(\mathbf{s} | I), \quad (8)$$

which is related to the Student t-distribution (Student, 1908).

If we were more knowledgeable, prior information regarding the source waveforms could be used to improve our inferences. In addition, knowledge of the source-detector coupling, which is found by solving the electromagnetic forward problem, could be utilized to create an algorithm that simultaneously performs source separation and localization (Knuth, 1998; Knuth & Vaughan, 1998). For simplicity, in this development we choose to assign uniform priors to $p(\mathbf{C} | I)$ and $p(\mathbf{s} | I)$, and absorb the terms into the implicit proportionality constant. It is more convenient to work with the logarithm of the posterior probability (8), which can be compactly written as

$$\ln P = -\frac{MRT}{2} \ln Q + \text{const}, \quad (9)$$

where P is the posterior probability $p(\mathbf{C}, \mathbf{s}(t), \boldsymbol{\alpha}, \boldsymbol{\tau} | \mathbf{x}(t), I)$.

An iterative algorithm to identify mcERPs, which we call *differentially variable component analysis* (dVCA), is implemented by solving (9) for the most probable set of model parameters. Such a solution is called the *Maximum A Posteriori* (MAP) estimate. This most probable set of model parameters is represented as a peak in the space of posterior probabilities. At this maximum, the partial derivative of the posterior probability with respect to any one of the model parameter values is zero. For this reason, our first step in deriving an optimal estimate of the component waveshape by equating each of the partial derivatives of (9) with zero. The partial derivative with respect to the j^{th} component waveshape at time q gives

$$\frac{\partial \ln P}{\partial s_j(q)} = -\frac{MRT}{2} Q^{-1} \frac{\partial Q}{\partial s_j(q)} \quad (10)$$

with

$$\frac{\partial \mathcal{Q}}{\partial s_j(q)} = -2 \sum_{m=1}^M \sum_{r=1}^R \left[W C_{mj} \alpha_{jr} - (C_{mj} \alpha_{jr})^2 s_j(q) \right] \quad (11)$$

where

$$W = x_{mr}(q + \tau_{jr}) - \sum_{\substack{n=1 \\ n \neq j}}^N C_{mn} \alpha_{nr} s_n(q - \tau_{nr} + \tau_{jr}). \quad (12)$$

The term W is important as it deals with the data, which has been time-shifted according to the latency shift of the component being estimated, $x_{mr}(q + \tau_{jr})$. From this, one subtracts all other components after each has been appropriately scaled and time-shifted, $C_{mn} \alpha_{nr} s_n(q - \tau_{nr} + \tau_{jr})$. The derivative of the log probability is zero when the scaled, estimated waveshape equals W . Thus one can obtain an expression for the optimal waveshape of the j^{th} source at time q in terms of the other sources

$$\hat{s}_j(q) = \frac{\sum_{m=1}^M \sum_{r=1}^R W C_{mj} \alpha_{jr}}{\sum_{m=1}^M \sum_{r=1}^R (C_{mj} \alpha_{jr})^2}. \quad (13)$$

Similarly for the source amplitudes, one obtains an optimal estimate by considering the derivative of the log probability with respect to the amplitude of the j^{th} source during the p^{th} trial. Setting this derivative equal to zero results in

$$\hat{\alpha}_{jp} = \frac{\sum_{m=1}^M \sum_{t=1}^T [U V]}{\sum_{m=1}^M \sum_{t=1}^T V^2} \quad (14)$$

where

$$U = \left(x_{mp}(t) - \sum_{\substack{n=1 \\ n \neq j}}^N C_{mn} \alpha_{np} s_n(t - \tau_{np}) \right) \quad (15)$$

and

$$V = C_{mj} s_j(t - \tau_{jp}), \quad (16)$$

such that the solution is given by the projection of the detector-scaled component $C_{mj} s_j(t - \tau_{jp})$ onto the data after removing the other scaled, time-shifted components. This can be viewed in terms of a dot product, which is related to a matching filter solution.

The optimal source-detector coupling coefficients are found similarly with

$$\hat{C}_{ij} = \frac{\sum_{r=1}^S \sum_{t=1}^T [X Y]}{\sum_{r=1}^S \sum_{t=1}^T Y^2} \quad (17)$$

where

$$X = \left(x_{ir}(t) - \sum_{\substack{n=1 \\ n \neq j}}^N C_{in} \alpha_{nr} s_n(t - \tau_{nr}) \right) \quad (18)$$

and

$$Y = \alpha_{jr} s_j(t - \tau_{jr}). \quad (19)$$

Estimating the latency shift of the j^{th} source during the p^{th} trial using the approach taken for the other parameters leads to a complex solution as the latency appears implicitly as the argument of the waveshape function. Instead we examine the necessary conditions for maximizing the quadratic form Q . Expanding the square in (7), one can see that as the latency shift τ_{jp} is varied, only the cross-terms corresponding to the j^{th} source change as long as the source waveshapes are zero outside of a closed time interval. The optimal estimate of the latency shift $\hat{\tau}_{jp}$ can be found by maximizing the cross-correlation between the estimated source and the data after the contributions from the other sources have been subtracted such that

$$\hat{\tau}_{jp} = \arg \max Z(\tau_{jp}) \quad (20)$$

where

$$Z(\tau_{jp}) = \sum_{m=1}^M \sum_{t=1}^T \left[C_{mj} \alpha_{jp} s_j(t - \tau_{jp}) \left(x_{mp}(t) - \sum_{\substack{n=1 \\ n \neq j}}^N C_{mn} \alpha_{np} s_n(t - \tau_{np}) \right) \right]. \quad (21)$$

In practice, as a discrete model is being used for the component waveshapes $\mathbf{s}(t)$, we utilize a discrete set of latency shifts with resolution equal to the sampling rate.

Perhaps the greatest challenge is determining the number of sources warranted by the data. In our investigations to date we have been focused on understanding the major sources responsible for the recorded data. This typically entails first modeling a single component and examining the un-modeled residual signal in detail. We then attempt to model the data with two components and so on to greater numbers of sources. As we will demonstrate, the responses we are examining are sufficiently complex that this approach rapidly reveals previously unknown characteristics of the neurophysiology. It is however, straightforward to apply more analytical methods of determining the appropriate numbers of sources. One such method is the Akaike Information Criterion (AIC) (Akaike, 1974; Victor & Canel, 1992). Increasing the complexity of the model by adding sources will always describe the data better and will increase the likelihood of the solution. However, by adding sources we increase the model complexity as we now have more model parameters. The idea is that a balancing point is reached when the increase

in likelihood is no longer worth the increase in model complexity. This point represents the optimal number of sources. For N sources in dVCA, the AIC is

$$\text{AIC}(N) = MRT \log(Q) + c^2 (NT + 2NR + N^2). \quad (22)$$

The first term is minus one times the logarithm of the likelihood. It is a data-dependent term, which decreases with an increasing number of sources due to the fact that more sources can better describe the data resulting in lower residual variance Q . The second term is a model-dependent term, which is a constant times the number of parameters in the model. This term increases with the number of sources. The constant c^2 is the t-statistic, with $c^2 = 4$ corresponding to a conservative $p = 0.95$. The solution for N sources is statistically significant ($p \leq 0.05$) when $\text{AIC}(N) < \text{AIC}(N - 1)$, which is equivalent to

$$\log(Q_{N-1}) - \log(Q_N) > 4 \frac{(T + 2R + 2N - 1)}{MRT}. \quad (23)$$

Note that Q_N and Q_{N-1} refer to the residuals of the model in (7) with N sources and $N - 1$ sources respectively. While this is an attractive technique, we have found that the number of sources warranted by this criterion is far greater than those we have been able to fully understand by analyzing the data manually source by source.

There are many ways that one can use the equations above to implement the dVCA algorithm. A useful iterative method (Figure 1) begins by modeling a single neural source with the single-trial amplitudes set to unity and the latency shifts set to zero.

1. Event-related potentials (ERPs) are computed for each detector by averaging the data over all trials. The ERP for each detector is full-wave rectified, and its integral (area under the curve) is approximated. The ERP having the largest area, or total signal content, is chosen as the initial approximation of the first component waveshape.
2. The single-trial amplitude scales are all initialized to one and the single-trial latency shifts are initialized to zero. This is consistent with the implicit assumptions of the ERP.
3. The coupling matrix is estimated using (17).
4. Single-trial latency shifts are estimated using (20).
5. Single-trial amplitudes are estimated using (14).
6. Equations (13), (17), (20), and (14) are then iterated until the average change in the waveshape from the previous iteration is less than 1% or until a maximum number of iterations has been performed.
7. At this point the residual signal for each detector is computed by subtracting the model from the data, as in the argument of (7). The residual signals are then averaged across trials to obtain a residual ERP for each channel.
8. If $N \geq 2$, the AIC criterion is applied. If satisfied (or if $N = 1$), another source is added and modeled.

9. The initial approximation of the next component waveshape is chosen to be the residual ERP of the detector having the largest total signal.
10. The single-trial amplitudes and latency shifts of the new component are set to one and zero, respectively.
11. Equations (13), (17), (20), and (14) are used to obtain the parameters accompanying the new component in addition to refining the estimate of the first component. The set of equations is further iterated to refine the solutions until the average changes in each component waveshape is again less than 1% or until the maximum number of iterations has been reached.
12. Additional components are added until the AIC criterion fails or the investigator chooses to stop.

This implementation represents only one possible approach to utilizing these equations to obtain useful solutions.

SIMULATIONS

The dVCA algorithm was evaluated using synthetic data to demonstrate the utility of amplitude and latency variability in the identification of multiple evoked components and also to assess the robustness of the algorithm in the presence of noise. We simulated electric field potentials recorded from a linear-array multielectrode with 15 channels spanning the cortical laminae in V1. Specifically, we designed three synthetic ERP components (Figure 2A) sampled at 2 kHz to approximate the neural ensemble response to diffuse red-light stimulation in macaque V1 (Givre et al., 1995; Mehta et al., 2000). Figure 2B shows the field potentials derived from the noise-free synthetic data as the detectors in the multielectrode would record them. Superimposing the field potentials from the three sources and approximating the second spatial derivative of this summed activity yields the current source density (CSD) profile shown in Figure 2C. The value of the CSD technique is evident as it both localizes the neural activity at the current sources and sinks, and eliminates volume-conducted activity (i.e. see c3 below).

The first component, c1, represents the initial biphasic activation in lamina 4, followed by the second component, c2, representing activation in the supragranular layers. The third component, c3, represents a far-field source that volume conducts to the electrode channels and is observable in the field potential waveforms (Figure 2B), but it is absent from the CSD plot (Figure 2C) as the coupling between this component and the channels is nearly constant (linear with a small slope). The spatial distribution of component amplitudes across the electrode array is defined by the coupling matrix, which was chosen to simulate the expected spatial distributions of the neural ensembles (current source-sink pairs) located in lamina 4, in the supragranular layers, and at a distant site. While this is a simplistic model, it captures features we expect to see in real recordings.

In the majority of simulations, uncorrelated Gaussian-distributed, additive noise was introduced, as in (1). The signal-to-noise ratio (SNR) is different for each component from trial to trial because the three simulated components have different single-trial amplitudes. For this reason, we specify the trial-average SNR for each component. As the mean amplitude scale of the components is set to one, this is easily computed from the standard deviation of the component waveform divided by the standard deviation of the additive noise,

$$\text{SNR}_{\text{component}} = 20 \log_{10} \frac{\sigma_{\text{component}}}{\sigma_{\text{noise}}} \quad (24)$$

where $\sigma_{\text{component}} = 0.357, 0.072$, and 0.260 for $c1$, $c2$, and $c3$ respectively. These σ values were calculated by taking the product of the standard deviation of the original component waveshapes and the absolute value of the maximum coupling coefficient for that source. Thus the SNRs reported in this paper represent the *maximal SNR* for each particular component.

The performance of the dVCA algorithm was evaluated in several ways. First, the ability of dVCA to separate the three mixed signals was measured using the Amari error (Amari et al., 1996), which derives from the fact that the estimated coupling matrix $\hat{\mathbf{C}}$ can be determined to within a scaled permutation of the true coupling matrix

$$\hat{\mathbf{C}} = \mathbf{C} \mathbf{\Sigma} \mathbf{\Pi}, \quad (25)$$

where $\mathbf{\Sigma}$ is a diagonal scaling matrix and $\mathbf{\Pi}$ is a permutation matrix. Here the quantities decorated with a carat denote estimated quantities and those without a carat denote the unknown true values. Ideally, the source estimates are therefore related to the original sources by a simple matrix transformation

$$\hat{\mathbf{s}} = \mathbf{M} \mathbf{s}, \quad (26)$$

where \mathbf{s} is the $N \times T$ matrix of the original source waveshapes, $\hat{\mathbf{s}}$ is the $N \times T$ matrix of the estimated source waveshapes, and \mathbf{M} is a transformation matrix found by

$$\mathbf{M} = \mathbf{\Pi}^{-1} \mathbf{\Sigma}^{-1}. \quad (27)$$

The deviation of \mathbf{M} from the ideal form suggested in (27) provides a measure of the quality of separation. We estimate \mathbf{M} from (26) by

$$\hat{\mathbf{M}} = (\hat{\mathbf{s}} \mathbf{s}^T)(\mathbf{s} \mathbf{s}^T)^{-1} \quad (28)$$

and compute the Amari error by summing the rescaled cross-terms, which describes the degree of component mixing

$$E_{\text{Amari}} = \frac{\sum_{i=1}^N \left(\sum_{j=1}^N \frac{|\hat{M}_{ij}|}{\max_k |\hat{M}_{ik}|} - 1 \right) + \sum_{j=1}^N \left(\sum_{i=1}^N \frac{|\hat{M}_{ij}|}{\max_k |\hat{M}_{kj}|} - 1 \right)}{2(N^2 - N)}. \quad (29)$$

Note that we have normalized the Amari error to have a maximum value of one rather than a number dependent on the number of sources.

Second, the ability of dVCA to estimate each component waveshape was evaluated by calculating the fractional RMS error of the estimated waveshape as compared to the correct waveshape. Note that before this comparison could be made the

estimated components $\hat{s}(t)$ needed to be rescaled and permuted because of the indeterminacy described by (25). For the j^{th} component

$$E_{wave}^j = \frac{\sqrt{\sum_{t=1}^T (s_j(t) - \hat{s}_j(t))^2}}{\sqrt{\sum_{t=1}^T (s_j(t))^2}}. \quad (30)$$

Last, the accuracies of the amplitude and onset latency estimates for the j^{th} component were evaluated by computing the deviation from the correct values for the entire set of single-trial estimates using

$$E_{amp}^{jr} = \alpha_{jr} - \hat{\alpha}_{jr} \quad (31)$$

and

$$E_{lat}^{jr} = \tau_{jr} - \hat{\tau}_{jr}, \quad (32)$$

and then computing the range of values contained within both the 68th and 95th percentiles. This not only provided a measure of accuracy, but also allowed us to compare the advantage of employing dVCA over the standard technique of averaging. This comparison is made by noting that with a trial-to-trial variability summarized by a standard deviation σ , the error one obtains by working with the average ERP is greater than or equal to the level of the variability σ . Thus errors in the dVCA estimates below the standard deviations of the variability represent a significant improvement over the standard assumption that amplitudes and latencies do not vary.

EFFECT OF AMPLITUDE VARIABILITY

We first demonstrated that amplitude variability aids in the separation process. Eleven experiments using synthetic data, each consisting of 50 simulated trials from the source component configuration described above, were performed where the degree of variability of the single-trial amplitudes was controlled. To generate the synthetic data, 50 single-trial amplitudes for each component were randomly sampled from a log-normal distribution with a sample mean $\mu_{amp} = 1.0$ and sample standard deviations of $\sigma_{amp} = \{0, 0.0625, 0.125, 0.1875, 0.219, 0.25, 0.375, 0.5, 0.625, 0.75, 1.0\}$ for each of the 11 sets of synthetic data. As latency variability was not simulated in this set of experiments, the single-trial latency parameters were set to zero ($\tau_{jr} = 0$). Given the component waveshapes, the coupling matrix, and the single-trial amplitudes and latencies, equation (1) was used to generate the synthetic data. A standard deviation of $\sigma_{noise} = 0.217$ was used for the noise resulting in SNRs of 4.3 dB, -9.6 dB, 1.5 dB (1.64, 0.33, 1.18 in terms of standard deviation ratios) for c1, c2, and c3 respectively. The dVCA algorithm was used to estimate the coupling matrix, component waveshapes, and single-trial amplitudes and latencies from the synthetic data. As the true parameters were known, the performance of the algorithm was evaluated as previously described.

Without amplitude variability, dVCA was unable to separate the components as demonstrated by the high Amari error of 0.219 (see Figure 3A). A small degree of

amplitude variability, $\sigma_{amp} = 0.25$, renders the problem solvable as demonstrated in Figure 3A, where the Amari error drops below 0.05 corresponding to about a 10% fractional RMS waveshape error (about a 23 dB SNR) for three equal variance components, and remains relatively constant with increasing variability hovering about an average of 0.028. Figure 3B shows the dramatic improvement in the waveshape estimation as quantified by the RMS errors, which rapidly drop to levels commensurate with the SNRs of the individual components. The effect of amplitude variability on the single-trial amplitude estimates (not shown) remained relatively constant for $\sigma_{amp} \geq 0.25$ with mean standard deviations of the errors in the single-trial amplitude estimates of 0.014, 0.076 and 0.010 and for c1, c2, and c3 respectively. This is well below the standard deviation of the amplitude variability. Last, even though the true onset latencies were set to zero in these simulations, dVCA still estimates these quantities. The errors of the onset latency estimates (not shown) also remained relatively constant with respective mean standard deviations of 0.417, 2.059, and 1.000 ms.

EFFECT OF LATENCY VARIABILITY

Next we demonstrated that latency variability also aids in the separation process. Eight experiments using synthetic data, each consisting of 50 synthetic trials from the source component configuration described above, were performed where the variability of the single-trial latencies was controlled. In these experiments there was no amplitude variability ($\alpha_{jr} = 1$). To generate the simulated data, 50 single-trial latencies were randomly drawn from a Gaussian distribution with sample mean $\mu_{lat} = 0$ and sample standard deviations of $\sigma_{lat} = \{0, 1.25, 2.5, 3.75, 5.0, 6.25, 7.5, 10.0\}$ ms for each of the eight simulations. The same noise variance was used as in the amplitude variability case.

Amari error was found to decrease with increasing latency variability, Figure 3C, dropping to below 0.05 with $\sigma_{lat} \geq 7.5$ ms. Component waveshape estimation was also found to improve with increasing onset latency variability, but the effect is not nearly as dramatic as in the amplitude variability case. Moreover, with the onset latency variability ranges considered, the accuracy of the algorithm's estimates never attained the levels found with amplitude variability. The most likely reason for this effect is that in terms of signal amplitude, amplitude variability is a first order effect whereas latency variability, when written as a Taylor expansion, is a second order effect dependent on the first derivative of the waveshape. While the amplitudes were all set to one in the simulated data, dVCA still estimates these quantities. Again, these amplitude estimate errors were low for $\sigma_{lat} \geq 7.5$ ms with the mean standard deviation of the errors in the estimates equal to 0.017, 0.077, 0.011 for c1, c2, and c3 respectively, which is on the order of the errors seen in the amplitude variability trials. Onset latency estimate errors also remained relatively constant with standard deviations of 0.250, 2.250, and 1.142 ms respectively.

SENSITIVITY TO ADDITIVE WHITE GAUSSIAN NOISE

In this first noise study we examined the robustness of dVCA in the face of additive white Gaussian noise by simulating twelve different noise levels. Each simulation relied on 50 trials of synthetic data where variability in both the amplitudes and latencies of the components was modeled. Variability ranges were in accordance with those observed in

our previous investigations (Truccolo et al. 2001). Single-trial amplitudes were drawn from a log-normal distribution with sample mean $\mu_{amp} = 1.0$ and sample standard deviation $\sigma_{amp} = 1.0$. Single-trial latencies were drawn from a normal distribution with sample mean $\mu_{lat} = 0$ and sample standard deviation $\sigma_{lat} = 10.0$ ms. The synthetic signal from each electrode (detector) was contaminated with a unique white Gaussian noise waveform as specified in (1).

The component specific SNRs and resulting Amari errors, listed in Table 1 and shown in Figure 4A, indicate a relatively smooth increase in Amari error with decrease in c1 SNR. However a significant jump in error occurs as the SNR of c2 passes below -25 dB (c1 SNR = -11 dB) suggesting that an SNR level of about -25 dB may denote a transition from a useful data set to a prohibitively noisy one. These results can be compared with the waveshape fractional RMS errors, which grow exponentially with decreasing SNR (Figure 4B). The errors reach 1.0, signifying that the deviations in the estimates are on the scale of the waveshape itself, at about -23 dB to -25 dB for the localized components and about -31 dB for the far-field component (c1 SNR = -28 dB). Figure 4E shows waveshape results for four different noise levels providing a better idea of the quality of the separation under these conditions. Most noteworthy is the fact that the majority of the waveshape error is due to the high-frequency contamination of the Gaussian noise rather than mixing of the components. Much of this could easily be improved by incorporating a prior probability describing the expectation that components should be slowly varying with respect to typical sampling rates, which would effectively filter the results. As we will demonstrate, explicitly filtering a real data set can remove important signals, and alter others (Mocks et al., 1986; Bogacz et al., 2002). Regardless, some distortion in the positive phase of c1 can be seen at a $SNR_1 = -10.8$ dB, which is much more easily noticeable at $SNR_1 = -22.8$ dB where only c1 and c3 remain detectable. The improved accuracy in the estimation of the far-field component c3 (illustrated in Fig. 4B) is most likely due to the fact that each electrode in the array provides information about the far-field source, whereas for local sources there are detectors with small signals.

Next we examine the quality of the single-trial amplitude estimates. Figure 4C shows the relationship between the errors of the estimates and the range of variability of the amplitudes. The performance of dVCA exceeds that of signal averaging when the estimation errors are less than the degree of variability in the original single-trial amplitudes. This is because by simply averaging the responses, one implicitly assumes that the single-trial amplitude is consistently equal to the mean resulting in errors equal to the degree of variability. The horizontal black lines in Figure 4C represent the degree of variability of the component amplitude as one standard deviation of the single-trial amplitudes about their mean. The solid blue and dashed red curves represent one and two standard deviations, respectively, of the single-trial amplitude estimate errors. Again, the far-field estimates (c3) were more accurate than those of the local sources (c1 and c2). Ninety-five percent of the amplitude estimates were within the degree of variability down to SNRs between -18 to -23 dB with 68% of the estimates within the range well down to -23 to -25 dB, which is consistent with the performance indicated by both the Amari error and the waveshape RMS error. To demonstrate the quality of the amplitude estimates, Figure 4F shows scatter plots of the true amplitude scales versus the estimates

for c1 at the same four SNR levels as in Figure 4E. Note that different values for the single-trial amplitudes were used in each simulation. Correct estimates will lie on the diagonal, whereas incorrect estimates are off diagonal. While the errors in the values of the amplitude estimates become unacceptable around -22.8 dB the pattern still exhibits a strong correlation ($r = 0.948$).

Figure 4D shows the behavior of the onset latency estimates, which are less well estimated than the amplitudes. In addition, the degradation of the quality of the far-field estimates was not noticeably different than those of the local sources with 95% of the estimates being within the range of variability down to an SNR of -8 to -12 dB, and 68% of the estimates within the range of variability down to -20 to -25 dB. Figure 4G shows scatter plots of the true onset latencies versus the estimates for c1. The diagonal pattern, which indicates a high level of predictability, is almost lost by -22.8 dB ($r = 0.392$). Note that due to the difference in variability levels of the amplitudes and onset latencies, the distribution of points in Figure 4F should not be directly compared to those in Figure 4G as they are effectively at different magnifications.

SENSITIVITY TO CORRELATED FAR-FIELD NOISE

In an effort to more accurately simulate the conditions that may be experienced during a real experiment, we designed a second noise study to determine the effect of ongoing correlated far-field activity on dVCA performance. The ongoing far-field activity was modeled as a time-series with a $1/f$ spectrum so that correlations would exist at all time scales. To simulate its far-field nature the ongoing activity was coupled identically to each detector.

Again twelve levels of the ongoing noise amplitude were tested; however, in this study Gaussian noise was not added to the individual electrodes. The specific noise levels, SNRs of the three components and resulting Amari errors are listed in Table 2. Figure 5A shows the Amari error smoothly increasing with decreasing SNR. The Amari error reaches the same level of error as in the Gaussian noise case with 20 dB less noise amounting to large errors around a c1 SNR of 0 to -5 dB. It is apparent that this type of noise more severely limits the ability of dVCA to separate signals. Figure 5B shows the component waveshape errors blowing up around -10 dB, with the effect on the far-field component c3 (red) being understandably catastrophic as both c3 and the noise are correlated across detectors. The amplitude errors reach unacceptable levels around -5 to -8 dB (Figure 5C). However, most interesting are errors of the onset latencies (Figure 5D). The errors for the local sources reach high levels (68% of estimates being within the range of variability) between -7 and -8 dB, whereas the errors for the far-field component c3 are large across the entire SNR range considered with 95% of the estimates never being within the range of variability. This is because the far-field noise is severely interfering with the ability to identify the far-field component in any given trial.

APPLICATION TO REAL DATA

To further demonstrate the utility of dVCA, we applied the algorithm to signals recorded from primary visual cortex of a male macaque monkey during the cuing period of an earlier attentional study (Mehta et al., 2000). During data collection, the subject was required to make a discrimination between standard and target visual stimuli for a juice reward. The standard visual stimulus was a $10\text{-}\mu\text{s}$ red-light flash presented through a

diffusing plate subtending a 20° visual angle, centered on the point of visual fixation, while the target was presented similarly but slightly differed in intensity. Stimuli were presented at irregular interstimulus (ISI) intervals (minimum of 350 ms, average of 2 stimuli per second). Intracortical field potentials were recorded using a linear-array multielectrode with 14 contacts, equally spaced at 150 μm , inserted into V1 and positioned so that the channels spanned all six laminae (see Figure 2A, middle). The continuous field potential record from all channels, incorporating the stimulus tags, was sampled at 2 kHz and recorded on a PC-based data acquisition system (Neuroscan, El Paso, Texas).

The signals examined with dVCA were recorded during 171 trial presentations of the standard visual stimulus and were epoched from 0 to 300 ms (0 ms indicates stimulus presentation). No other pre-processing of the data was performed. The average field potential signals in each electrode contact were calculated and utilized to determine the current source density (CSD) profile of these data (Figure 6A). The CSD profile was approximated using a 3-point, second-order difference of the field potentials (Nicholson & Freeman, 1975; Schroeder et al., 1995). This profile is a useful tool because it indexes the local regions of transmembrane current flow and eliminates volume-conducted activity generated at a distant site, which often contaminates local field potential recordings. Consistent with earlier studies reviewed by Schroeder et al. (1995; 2001), the visually-evoked CSD profile sampled from V1 during this experimental session shows that the earliest activation occurs in the granular subdivisions of Layer 4, which are the major target of thalamic afferents. A second focus of activity localizes to the supragranular (laminae 2 and 3) layers and is thought to index a feedforward activation of pyramidal cells by interneurons in the granular layer (Schroeder et al. 1990; 1991). Note that as we are using a 3-point, second-order difference to approximate the CSD, the top waveform in Figure 6A is associated with electrode channel 2 rather than channel 1.

Although it is an attractive idea to apply dVCA as an automated computer program that analyzes the data and produces an answer (as we did with simulated data), we demonstrate that it must be used manually as a tool to understand the data being analyzed. To stress the importance of this, instead of following the flow chart exactly as in Figure 1, we first applied dVCA to the single-trial local field potential signals to extract a single component. As shown in Figure 6B, the CSD profile of this component illustrates a biphasic response localized in the granular layer suggesting that it represents the initial response to the thalamic inputs. Not surprisingly, this waveshape and its distribution are very similar to the CSD profile of the ensemble averaged ERP of Figure 6A, as the initial approximation to the first component derives from the ensemble average. However, note that it excludes much of the minor variations not associated with this main activation pattern. In addition to providing information about the component waveshape and its spatial distribution, dVCA provides information about the amplitude and latency shift of the component in each trial. Figure 6C shows the distribution of single-trial amplitude scales for this component. In accordance with the dVCA normalization condition, the mean of the sample equals one ($\mu_1 = 1$). More importantly, the distribution illustrates that there is very little single-trial amplitude variability ($\sigma_1 = 0.1044$), which is on the order of 10% variation. In contrast, the distribution of single-trial latency shifts (Figure 6D) shows something quite unexpected – it is bimodal with an early latency peak at -4.625 ms and a late latency peak at 2.125 ms with a

difference of 6.75 ms. Moreover, the ratio of late to early responses is 2.29:1 (119:52), as defined by a -1.25 ms cutoff between the two modes.

We performed several different analyses to confirm the existence of the bimodal latency distribution displayed in Figure 6D. First, Figure 6E displays a colorized plot of the actual, single-trial field potentials (FPs) recorded from electrode channel 10, which is located in the granular layer. Time runs along the horizontal axis, chronological trial number runs along the vertical axis, and color indicates the amplitude of the FP. The red dashes on the left side of the plot indicate trials for which dVCA determined the latency to be early. It is easily seen that these highlighted trials have a response onset before the late trials, and it confirms that the bimodal latency distribution of the dVCA estimation concurs with observations from the actual data. Second, we performed a cross-correlation analysis between each single-trial waveshape and the trial-averaged, FP waveform between 25 and 95 milliseconds (this time period captures the initial negative deflection of the averaged FP waveform). This analysis also yielded a bimodal distribution (not shown). Finally, we selectively averaged the early and late FP recordings in channel 10 and showed them to be significantly different in both latency and waveshape (Figure 6F). The trial-averaged FP waveshape (black) lies in-between the early (red) and late (blue) waveshapes. As expected, the minimum of the early averaged responses occurs before the minimum of the late averaged responses (5.5 ms difference). While this time difference reflects the fact that these early responses precede the late responses, it is however not as accurate as the dVCA estimate as it is derived from a single-time point rather than from the entire waveshape. It is also not surprising that the late responses more closely resemble the ensemble averaged FP as they outnumber the early responses by more than 2 to 1.

Figure 6E illustrates one more interesting phenomenon: the early and late responses tend to be grouped in chronologically in time, indicating that the effect is not random. This suggests that the two response modes may be related to the attentional or arousal state of the subject.

As these two types of granular responses suggest different physiological mechanisms, we split the data set into two subsets: an early subset and a late subset. We then used dVCA to re-estimate the first component in each subset. Figures 7A and B show the CSD of the first component from the early and late subsets, respectively. The onset latency of the prominent granular sink over source configuration was determined by descending down the left side of the peak and identifying the point in time where five consecutive previous data points were monotonically increasing. With this measure the response onset latency of the early subset (Figure 7A) was 28 ms, which is clearly earlier than its counterpart in the late subset (Figure 7B), which was 42.5 ms, as indicated by the drop lines. The waveshapes of the two types of Layer 4 responses differ significantly indicating a difference in activation patterns between subsets. This is also confirmed by the different degrees to which this activation appears in the supragranular layers, as the later responses seem to be more strongly coupled to the supragranular activation than the early responses.

After any analysis the residual signals must be examined since they represent activity that was not modeled by the algorithm. In Figure 7C we show the trial-averaged FP residuals for the early (red) and late (blue) subsets for odd-numbered electrode channels. First, the black arrows indicate activation that is almost identical in both

subsets (mean peak time at 37.75 ms for the early subset and 37.68 ms for the late subset). Due to the timing and the laminar distribution of this field potential, we believe that this negativity reflects the initial signal from the thalamocortical afferents. The early responses onset just after onset of this thalamic signal, which occurs at about 25 ms, whereas the late responses are delayed. After 50 ms the residual signals from the two subsets diverge significantly. This difference is striking in channel 1 where the oscillations are 180 degrees out of phase. Of significant interest are the ultra-high frequency (160-220 Hz) oscillations (UHF) in the granular and the supragranular layers of the early subset (red arrow in channel 11). These UHF oscillations, which onset with the initial granular response, are absent in the averaged residuals of late responses.

To verify that these UHF oscillations were not present in the late responses, we applied a Morlet-based wavelet transformation to the two residual subsets of channel 11. To preserve information about time-frequency behavior of oscillations not precisely time-locked to the stimulus, we applied the wavelet transform to each single-trial separately and the wavelet results were then averaged. These results show that there is a burst of UHF activity ranging from about 160-220 Hz and occurring between 42-57 ms in the early subset, which is represented as an “island of activity” in the time-frequency plot (Figure 7D). Such an island of activity is completely absent in the late subset (Figure 7E). This finding demonstrates the ability of dVCA to identify distinct physiological modes of activation using the single-trial characteristics of the responses, even in situations where the data are unfiltered.

To demonstrate how dVCA can be used to extract multiple components, we will now focus on the late subset. A complete analysis of the entire data set will be published elsewhere. In this example we model three components (Figures 8A, B, C). Components 2 and 3, (c2 and c3) have been magnified by a factor of 2 to make them easily visible, and are thus not on the same scale as component 1 (c1), which is the largest response. To assure that the SNR is within the acceptable range for dVCA, we used equation (24) to compute the SNR for each component in each channel. For a given component this is accomplished by estimating the standard deviation of the component in that channel and the standard deviation of the residual signal in that channel. The average SNR for each component was then computed by averaging its SNRs across channels. We found that all three components are well within the range of acceptable performance with average SNRs of 1.59 dB, 1.68 dB, 0.98 dB for c1, c2, and c3 respectively.

First, c1 is in all practical respects identical to the single component we previously estimated from the late subset. The prominent granular sink over source configuration is activated at about 35.5 ms in this more detailed model. While the onset time is easy to quantify in the large biphasic response, the more complex oscillatory components required a more sophisticated onset measure. By fitting the pre-response interval with a line, we computed the standard deviation of the pre-response signal deviations from that line. Component onset was then defined as the first of 5 consecutive time points where the component amplitude was greater than two standard deviations. This was useful as we did not filter the data and low-frequency oscillations could produce confounding baseline deviations. With this measure, we found small, but significant, activations in c1 as early as 26 ms (see drop line in Fig 8A) occurring after the thalamic signal at 25 ms and before the massive biphasic response at 35.5 ms.

A low frequency biphasic response is described by c2 (Fig 8B). Significant onset of this component occurs at 37 ms (see drop line), which is after c1 begins the biphasic activation in the granular layers. Note that the source and sink distribution of c2 is in superficial layers as compared to the small supragranular response of c1. At first glance it appears that the spatial distribution of c3 is the same as that of c2. However, while these components may involve the same populations of cells, c3 also involves cells in the granular layers. The initial activation of c3 is at 30 ms, but it is not until 45 ms when the component begins to display a biphasic activation followed later by a slow wave.

Figure 8D shows the single-trial amplitude histograms along the diagonal, along with scatter plots showing the relationships between the single-trial amplitudes of the three components. The initial granular response, c1, has the lowest amplitude variability ($\sigma_{amp1} = 0.102$), whereas the later supragranular responses display greater variability ($\sigma_{amp2} = 0.248, \sigma_{amp3} = 0.362$). Figure 8E shows the single-trial latency results along with scatter plots depicting the trial-by-trial relationships between component latencies. Component c1 again displays the least variability ($\sigma_{lat1} = 1.08$ ms), but in this case, c2 displays greater variability than c3 ($\sigma_{lat2} = 9.85$ ms, $\sigma_{lat3} = 1.58$ ms).

Correlations in the scatter plots between component parameters indicate dynamic interactions among these components. Of the amplitude-amplitude interactions (Figure 8D), we see that the amplitudes of c2 are correlated with c3 ($r = 0.495, p < 10^{-7}$), so that when c2 is larger than average, c3 is also larger than average. Such correlation might suggest that these components have not been separated. However, note that the latency variability of each component is very different (Figure 8E), and that there is little correlation between the latency of c2 and the latency of c3 ($r = 0.171, p = 0.075$).

Of the latency-latency interactions, the largest correlation is between the c1 and c2 latencies ($r = 0.327, p < 0.001$), so that when c1 is earlier than average, c2 is also earlier than average. More interesting are the amplitude-latency interactions. Figure 8F shows the two most probable relationships. First the amplitude and latency of c1 are correlated ($r = 0.297, p < 0.002$), so that when c1 is early its response is smaller, and when it is late its response is larger. Such a self-interaction contains much information about the underlying dynamics of the neural response represented by c1. More apparent is the relationship between the amplitude of c1 and the latency of c2 ($r = 0.483, p < 10^{-7}$). In this case, when the c1 granular response is larger than average, c2 onsets later than average. This result is somewhat counterintuitive, as generally we expect that if the c1 is driving the c2, a larger c1 might produce an earlier onset c2. The difference between c1 and c2 is further highlighted by noting that the amplitude of c1 is anti-correlated to the latency of c3 ($r = -0.190, p = 0.048$) (not shown).

DISCUSSION

In this paper, our goal was not to understand the details of these responses in visual cortex and their dynamic interactions, but merely to describe how techniques like dVCA now enable neuroscientists to tease apart and study these dynamic interactions among neural ensembles during cognitive or sensorimotor processing. Detailed studies of the single-trial interactions among neural components in a variety of experimental situations

will provide much insight into the information processing strategies employed by the various cortical areas and by the brain as a whole.

Maximum likelihood techniques have been used previously to approach the problem of trial-to-trial variability of evoked responses. These past works have relied on single-component signal models that incorporate latency variability (Woody, 1967; Pham et al., 1987), and more recently both amplitude and latency variability (Jaskowski & Verleger, 1999) to characterize evoked responses. Multiple-component models were introduced by Lange et al. (1997) by adopting a template model of the ERP waveshapes, and were used to characterize both the amplitude and latency of multiple components in single-channel EEG recordings. Early versions of the dVCA algorithm also dealt with multiple-component, single-channel estimation (Knuth et al., 2001; Truccolo et al., 2001; 2002; 2003), but instead employed a completely general waveshape model that did not rely on a given functional form. The dVCA algorithm presented here allows one to use multiple channel recordings to identify and characterize multiple ERP components in the single trial by taking advantage of the trial-to-trial variability.

The dVCA algorithm was derived by approaching the problem as an exercise in Bayesian parameter estimation. There are several advantages to a Bayesian approach. First, the strategy is model-based in the sense that given a quantitative model of the phenomena of interest, probability theory can be used to estimate the values of the model parameters from the data. Any failure in the algorithm can be traced back to either inadequacies of the model to represent the physical situation, assumptions or approximations made in its implementation, or a situation produced by an insufficiently informative data set. Second, any inadequacies in the model, once identified, are readily remedied leading to improvements in the algorithm. This is the crux of the scientific method. Third, once the model component set has been adequately modeled, the residual data can be examined to identify previously hidden phenomena. For example, in this paper we have discovered UHF oscillations associated with early initial granular layer responses and absent in the more frequent late responses. By accurately identifying the evoked components in the single-trial epochs, one can more accurately study the ongoing activity, which has been purported to contain signals important for communication among brain regions (Bressler et al., 1993; Singer, 1993; Truccolo et al., 2003), perception, working memory, and sensorimotor integration (Engel et al., 2001; Lee et al., 2003). Fourth, the Bayesian methodology allows one to incorporate additional prior information into the problem to improve one's inferences, which is a major advantage that we plan to capitalize on in future work.

The dVCA algorithm itself has a number of technical strengths. The first is that neither the components, nor their underlying neural sources are assumed to be independent of one another. This avoids the adoption of physiologically implausible assumptions and enables one to study the dynamical interactions among neural sources. The second strength is that by taking into account the trial-to-trial variability in amplitude and latency, one is able to quantify the interactions among sources as well as study their time-dependent properties over the duration of the entire experiment. Single-trial analysis also makes it possible for dVCA to detect components, which are not present in every experimental trial thus allowing an experimenter to study cognitive or sensorimotor processes, which may employ different processing strategies at different times. Additionally, one can use dVCA to study how attention, arousal, and varying disease

states may modify neural responses on a single-trial basis. Finally, while no explicit prior information about the values of the model parameters was included in the development of the algorithm, much prior information went into the choice of the model. This is in contrast to other approaches such as ICA and other general BSS techniques, which strive to make very general assumptions about the model and the distributions of the parameter values (Knuth, 1997; 1999).

We have found that dVCA is robust in the presence of noise allowing accurate estimates of all parameters down to SNRs on the order of -20 dB for white Gaussian noise and -7 dB for correlated far-field noise. The estimation of the far-field signals in the presence of correlated far-field noise was difficult. This is undoubtedly due to the fact that the noise in this case possesses the same spatial distribution as the source making the two difficult to distinguish. In reality however, the behavior of the background noise will fall somewhere between these two extremes. When using dVCA, we recommend that one calculate the SNR of the estimated components to assure that the algorithm is operating in a regime where the quality of the parameter estimates is assured. In addition, the spatial distribution of the sources across the array should be examined to assure that they are physiologically reasonable. We have been able to construct cases, which remain inseparable by the algorithm. This can typically be detected by examining the CSD map of the estimated components across the array. In cases where two sources remain mixed, each are characterized by multiple sets of neural sources in identical locations. However, we have found this to be more difficult to deal with in intracortical laminar recordings as the cortical architecture only allows for a small number of current sources and sinks, and tight couplings between sources is extremely likely (Shah et al, 2002). One possible solution to this problem is to restart the algorithm using starting conditions consistent with the source locations indicated by the CSD. The value of the posterior probability of these solutions obtained from different starting points can also be computed and compared (when using the same model order) to assure that one has the most probable solution and is not merely stuck at a local maximum. Searching such enormous solution spaces is a difficult problem faced by all source separation algorithms.

We are currently working to improve the dVCA algorithm along several lines. First, there are often situations where the experimenter has knowledge about the forward problem, which describes the propagation of the signals to the detectors. Such knowledge can be incorporated by adopting a more specific source model (eg. current dipole model) and expressing the coupling coefficients in terms of the new ERP source parameters, the detector coordinates, and head geometry (Knuth & Vaughan, 1998), or by using information about the propagation law to derive appropriate prior probabilities for the coupling matrix elements (Knuth, 1998; 1999). Similarly, more detailed information about the correlation structure of the background noise can be used to derive more accurate likelihood functions (Sivia, 2003). Second, by employing a discrete model of the component waveshapes, we are restricted to estimating only discrete values of onset latency shift. By adopting a waveshape model that relies on a set of continuous basis functions, continuous values of latency shift could be investigated. An example of a continuous time model is the frequency domain model of the ERP employed by Pham et al. (1987) and Jaskowski & Verleger (1999). In that model the waveshape is described as a sum of a discrete set of sinusoids, which is continuous in the time domain, but discrete in the frequency domain. In principle, such a model allows one to describe latency shifts

with arbitrary precision. Third, this algorithm represents a MAP estimate based on an iterative, or fixed-point solution. While each step in this algorithm has intuitive appeal, it is perhaps not the most effective means to obtaining a solution. Markov chain Monte Carlo with simulated annealing is well suited to performing simultaneous model selection calculations and parameter estimations. We have briefly investigated such an approach only to find it to be extremely time consuming given the large number of parameters in the model. It is expected that a clever implementation based on the mcERP model and posterior probability derived in this paper would outperform the algorithm presented here by automatically identifying the number of components warranted by the data, avoiding local optimal solutions, and readily providing error bars for the results. Last, in principle this algorithm is equally applicable to human scalp EEG and MEG data. In practice, there are several challenges. The number of detectors utilized in these paradigms is typically an order of magnitude greater than those we have demonstrated here. It is certain that the algorithm in its present implementation will run more slowly. Also, whole-head studies expose the experimenter to an order of magnitude more sources than we work with in our intracortical recordings. This increases the possibility that the dVCA algorithm can become trapped in non-optimal local solutions. This of course is a potential problem for all source separation and localization techniques, and dVCA is no exception. We are beginning to examine the application of dVCA in human scalp studies, and expect that if such pathological solutions are encountered, they can be avoided by utilizing more sophisticated search algorithms.

ACKNOWLEDGEMENTS

KHK supported in part by a NARSAD Young Investigator Award, the NASA IDU/IS/CICT Program and the NASA Aerospace Technology Enterprise. AS supported by the Medical Scientist Training Program (T32M07288). CES supported by NIMH MH060358. SLB and MD supported by grants from NSF (IBN0090717) and NIMH (MH64204 and MH42900). We thank Drs. Ashesh Mehta and Istvan Ulbert for data collection, Dr. Peter Lakatos for helpful comments and discussion, and Dr. Len Trejo and Sam Clanton of NASA Ames Research Center for their assistance in optimizing the coding of portions of the algorithm.

REFERENCES

- Akaike H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Cont.* 19(6), 716-723.
- Amari, S., Cichocki, A. and Yang, H.H. (1996). A new learning algorithm for blind signal separation. In *Advances in Neural Information Processing Systems 8*. D. Touretzky, M. Mozer, and M. Hasselmo, eds. (Cambridge, MA: MIT Press), pp. 752-763.
- Bell, A.J., and Sejnowski, T.J., (1995). An information-maximization approach to blind source separation and deconvolution. *Neural Comp* 7, 1129-1159.
- Belouchrani, A., Abed-Meraim, K., Cardoso J.-F., and Moulines, E. (1993). Second-order blind separation of correlated sources. In *Proc. Int. Conf. on Digital Sig. Proc.*, Nicosia, Cyprus, pp. 346-351.
- Bogacz, R., Yeung, N., Holroyd, C.B. (2002) Detection of phase resetting in the electroencephalogram: an evaluation of methods, *Soc. Neurosci. Abstr.*, Vol. 28, Prog. No. 506.9
- Bressler, S.L., Coppola, R., Nakamura, R. (1993). Episodic multiregional cortical coherence at multiple frequencies during visual task performance. *Nature* 366, 153-156.
- Cao, J., Murata, N., Amari, S., Cichocki, A., Takeda, T., Endo, H., and Harada, N. (2000). Single-trial magnetoencephalographic data decomposition and localization based on independent component analysis approach. *IEICE T Fund Electr* 9, 1757-1766.
- Engel AK, Fries P, Singer W (2001). Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2, 704-716.
- Givre, S.J., Arezzo, J.C., Schroeder, C.E. (1995). Effects of wavelength on the timing and laminar distribution of illuminance-evoked activity in macaque V1. *Vis Neurosci* 12(2), 229-39.
- Hyvärinen, A., and Oja, E. (1997). A fast fixed-point algorithm for independent component analysis. *Neural Comp* 9, 1483-1492.
- Jaskowski, P., Verleger, R., (1999). Amplitude and latencies of single-trial ERP's estimated by a maximum-likelihood method. *IEEE Trans. Biomed. Eng.* 46, 987-93.
- Jaynes, E.T. (2003). *Probability Theory - The Logic of Science*. (Cambridge: Cambridge University Press).
- Jung, T.-P., Makeig S., Westerfeld M., Townsend J., Courchesne E., Sejnowski T.J. (1999). Independent component analysis of single-trial event-related potentials. In *Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation: ICA'99*. J. F. Cardoso, Ch. Jutten, Ph. Loubaton, eds. (Aussois France: ICA'99), pp. 173-8.
- Knuth, K.H. (1997). Difficulties applying recent blind source separation techniques to EEG and MEG. In *Maximum Entropy and Bayesian Methods*, Boise 1997, G.J. Erickson, J.T. Rychert and C.R. Smith, eds. (Dordrecht: Kluwer Academic Publishers), pp. 209-222.
- Knuth, K.H. (1998). Bayesian source separation and localization. In *Proceedings of SPIE: Bayesian Inference for Inverse Problems*, vol. 3459, A. Mohammad-Djafari, ed. (Bellingham WA: SPIE), pp. 147-58.

- Knuth, K.H. (1999). A Bayesian approach to source separation. In Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation: ICA'99. J. F. Cardoso, Ch. Jutten, Ph. Loubaton, eds. (Aussois, France: ICA'99), pp. 283-288.
- Knuth, K.H., Vaughan, Jr., H.G. (1998). Convergent Bayesian formulations of blind source separation and electromagnetic source estimation. In *Maximum Entropy and Bayesian Methods, Garching, Germany 1998*, W. von der Linden, V. Dose, R. Fischer, R. Preuss, eds. (Dordrecht: Kluwer Academic Publishers), pp. 217-26.
- Knuth, K.H., Truccolo, W.A., Bressler, S.L., Ding, M. (2001). Separation of multiple evoked responses using differential amplitude and latency variability. In Proceedings of the Third International Workshop on Independent Component Analysis and Blind Signal Separation: ICA 2001. T.-W. Lee, T.-P. Jung, S. Makeig, T.J. Sejnowski, eds. (San Diego, CA: ICA 2001).
- Lange, D.H., Pratt, H., Ingbar, G.F. (1997). Modeling and estimation of single evoked brain potential components. *IEEE Trans. Biomed. Eng.* 44, 791-9.
- Lee, K.H., Williams, L.M., Breakspear, M., Gordon, E. (2003). Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia. *Brain Res Brain Res Rev* 41, 57-78.
- Makeig, S., Jung, T.-P., Bell, A., Ghahremani, D., Sejnowski, T.J. (1997). Blind separation of auditory event-related brain responses in independent components. *Proc. Natl. Acad. Sci. USA* 94, 10979-84.
- Makeig, S., Westerfield, M., Jung, T.-P., Enghoff, S., Townsend, J., Courchesne, E., Sejnowski, T.J. (2002). Dynamic brain sources of visual evoked responses. *Science* 295, 690-694.
- Mehta, A.D., Ulbert, I., Schroeder, C.E. (2000). Intermodal selective attention in monkeys. I: distribution and timing of effects across visual areas. *Cereb Cortex* 10(4), 343-58.
- Mocks, J., Gasser, T., Kohler, W., De Weerd, J.P. (1986). Does filtering and smoothing of average evoked potentials really pay? A statistical comparison. *Electroencephalogr Clin Neurophysiol* 64, 469-480.
- Nicholson C., Freeman J.A. (1975). Theory of current source-density analysis and determination of conductivity tensor for anuran cerebellum. *J. Neurophysiol.* 38(2), 356-368.
- Pham, D.T., Mocks, J., Kohler, W., Gasser, T. (1987). Variable latencies of noisy signals: estimation and testing in brain potential data. *Biometrika* 74, 525-33.
- Särelä, J., Vigário, R., Jousmäki, V., Hari, R., Oja, E. (1998). ICA for the extraction of auditory evoked fields. In 4th International Conference on Functional Mapping of the Human Brain. (Montreal Canada: HBM'98).
- Schroeder, C.E. (1995). Defining the neural bases of visual selective attention: Conceptual and empirical issues. *Int J Neurosci* 80, 65-78.
- Schroeder, C.E., Tenke, C.E., Givre, S.J., Arezzo, J.C., Vaughan, H.G., Jr. (1990). Laminar analysis of bicuculline-induced epileptiform activity in area 17 of the awake macaque. *Brain Res.* 515(1-2), 326-30.
- Schroeder, C.E., Tenke, C.E., Givre, S.J., Arezzo, J.C., Vaughan, H.G., Jr. (1991). Striate cortical contribution to the surface-recorded pattern-reversal VEP in the alert monkey. *Vision Res.* 31, 1143-57.

- Schroeder, C.E., Steinschneider, M., Javitt, D.C., Tenke, C.E., Givre, S.J., Mehta, A.D., Simpson, G.V., Arezzo, J.C., Vaughan, H.G., Jr. (1995). Localization of ERP generators and identification of underlying neural processes. *Electroencephalogr Clin Neurophysiol Suppl* 44, 55-75.
- Schroeder, C.E., Mehta, A.D., Givre, S.J. (1998). A spatiotemporal profile of visual system activation revealed by current source density analysis in the awake macaque. *Cereb Cortex* 8(7), 575-92.
- Schroeder, C.E., Mehta, A.D., Foxe, J.J. (2001). Determinants and mechanisms of attentional modulation of neural processing. *Front Biosci.* 6, D672-84.
- Shah A.S., Knuth K.H., Truccolo W.A., Ding M., Bressler S.L., Schroeder C.E. (2002). A Bayesian approach to estimating coupling between neural components: evaluation of the multiple component event related potential (mcERP) algorithm. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Moscow ID 2002, AIP Conference Proceedings 659, C. Williams, ed., (Melville, NY: American Institute of Physics), pp. 23-38.
- Singer W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annu Rev Physiol* 55, 349-374.
- Sivia, D.S. (1996). *Data Analysis. A Bayesian Tutorial* (Oxford: Clarendon Press).
- Sivia, D.S. (2003). Some thoughts on correlated noise. In Press: *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Jackson Hole WY 2003, AIP Conference Proceedings, G.J. Erickson, K.H. Knuth and C.R. Smith, (eds.), (Melville, NY: American Institute of Physics).
- Student. (1908). The probable error of a mean. *Biometrika* 6, 1-24.
- Tang, A.C., Pearlmutter, B.A., Malaszenko, N.A., Phung, D.B. (2002). Independent components of magnetoencephalography: single-trial response onset times. *NeuroImage* 17, 1773-1789.
- Truccolo, W.A., Knuth, K.H., Ding, M., Bressler, S.L. (2001). Bayesian estimation of amplitude, latency and waveform of single trial cortical evoked components. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, Baltimore MD 2001, AIP Conference Proceedings 617, R.L. Fry and M. Bierbaum, eds. (Melville, NY: American Institute of Physics), pp. 64-73.
- Truccolo, W.A., Ding, M., Knuth, K.H., Nakamura, R., Bressler, S.L. (2002). Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity. *Clin Neurophysiol* 113(2), 206-26.
- Truccolo, W.A., Knuth, K.H., Shah, A.S., Bressler, S.L., Schroeder, C.E., Ding, M., (2003), Estimation of single-trial multi-component ERPs: Bayesian foundation and applications. In press: *Biol Cybern.*
- Victor, J.D., Canel A. (1992). A relation between the Akaike criterion and reliability of parameter estimates, with application to nonlinear autoregressive modeling of ictal EEG. *Annals of Biomedical Engineering* 20, 167-180.
- Vigário, R., Särelä, J., Jousmäki, V., Oja, E. (1999). Independent component analysis in decomposition of auditory and somatosensory evoked fields. In *Proceedings of the First International Workshop on Independent Component Analysis and Signal Separation: ICA'99*. J. F. Cardoso, Ch. Jutten, Ph. Loubaton, eds. (Aussois, France: ICA'99), pp. 167-72.

Woody, C.D. (1967). Characterization of an adaptive filter for the analysis of variable latency neuroelectric signals. *Medical Biological Engineering* 5, 539-53.

Additive White Gaussian Noise					
Case	Noise Std Dev	Component SNR (dB)			Amari Error
		1	2	3	
1	0.155	7.2	-6.7	4.5	0.004
2	0.310	1.2	-12.7	-1.5	0.005
3	0.437	-1.8	-15.7	-4.5	0.013
4	0.618	-4.8	-18.7	-7.5	0.011
5	0.734	-6.3	-20.2	-9.0	0.014
6	0.873	-7.8	-21.7	-10.5	0.026
7	1.037	-9.3	-23.2	-12.0	0.017
8	1.233	-10.8	-24.7	-13.5	0.050
9	1.742	-13.8	-27.7	-16.5	0.108
10	2.460	-16.8	-30.7	-19.5	0.100
11	4.909	-22.8	-36.7	-25.5	0.198
12	9.794	-28.8	-42.7	-31.5	0.421

Table 1. This table shows the effect of additive white Gaussian noise on the separability of the three components. The component SNRs are calculated using the component standard deviations of 0.357, 0.072, 0.260 for components c1, c2, and c3 respectively.

Additive 1/f Far-Field Noise					
Case	Noise Std Dev	Component SNR (dB)			Amari Error
		1	2	3	
1	0.036	20	6.1	17.2	0.015
2	0.048	17.5	3.6	14.7	0.013
3	0.063	15	1.1	12.2	0.017
4	0.113	10	-3.9	7.2	0.008
5	0.150	7.5	-6.4	4.7	0.035
6	0.201	5	-8.9	2.2	0.068
7	0.268	2.5	-11.4	-0.3	0.113
8	0.357	0	-13.9	-2.8	0.143
9	0.476	-2.5	-16.4	-5.3	0.159
10	0.634	-5	-18.9	-7.8	0.191
11	0.846	-7.5	-21.4	-10.3	0.359
12	1.128	-10	-23.9	-12.8	0.365

Table 2. This table shows the effect of additive 1/f correlated far-field background signal on the separability of the three components. The component SNRs are calculated using the component standard deviations of 0.357, 0.072, 0.260 for components c1, c2, and c3 respectively.

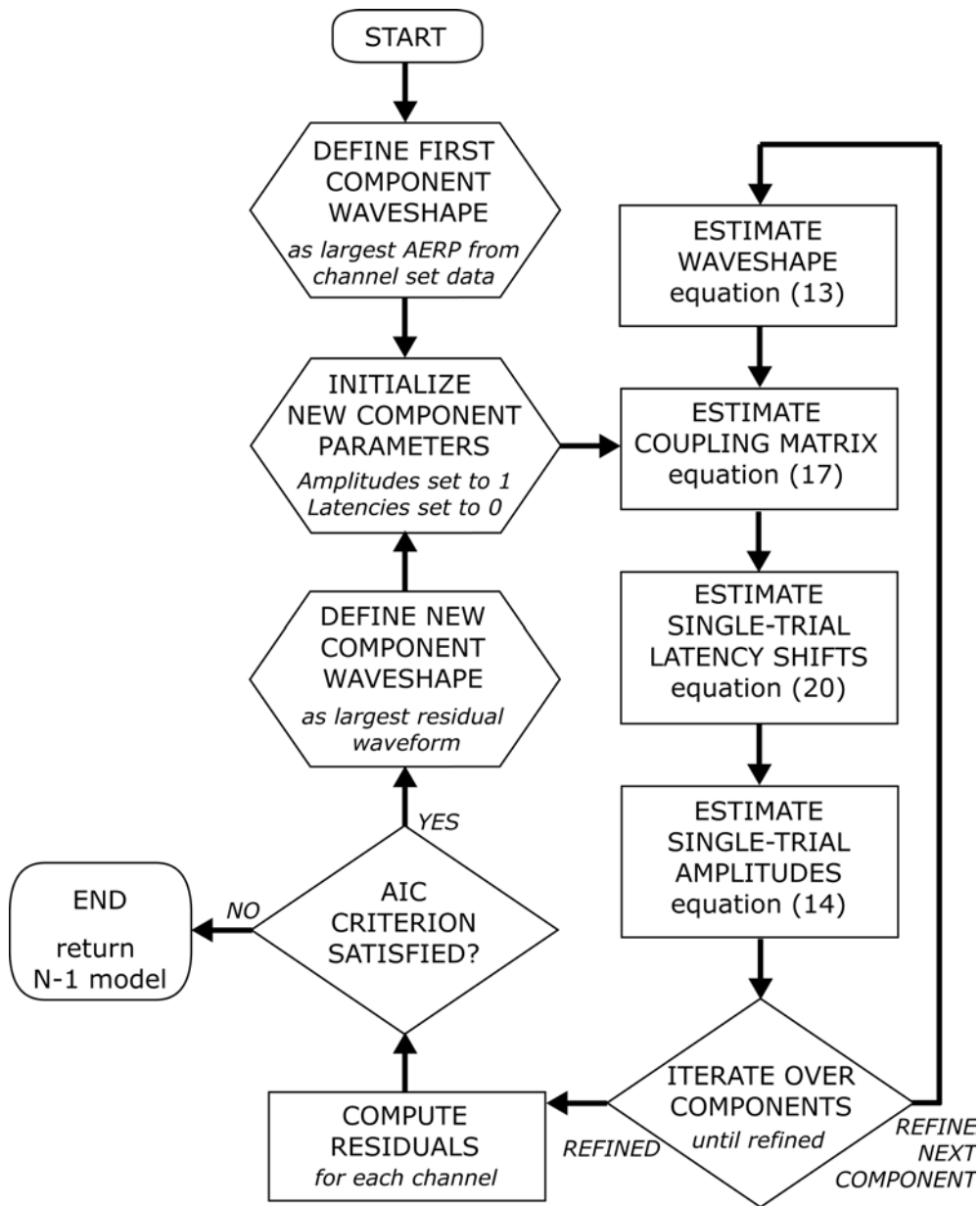


Figure 1 Flow chart describing the dVCA algorithm.

This flow chart describes the implementation of the dVCA algorithm, which relies on equations (13), (17), (20), and (14) in the text. It is based on a hierarchical model, which begins with a single component model and adds components until AIC is no longer satisfied.

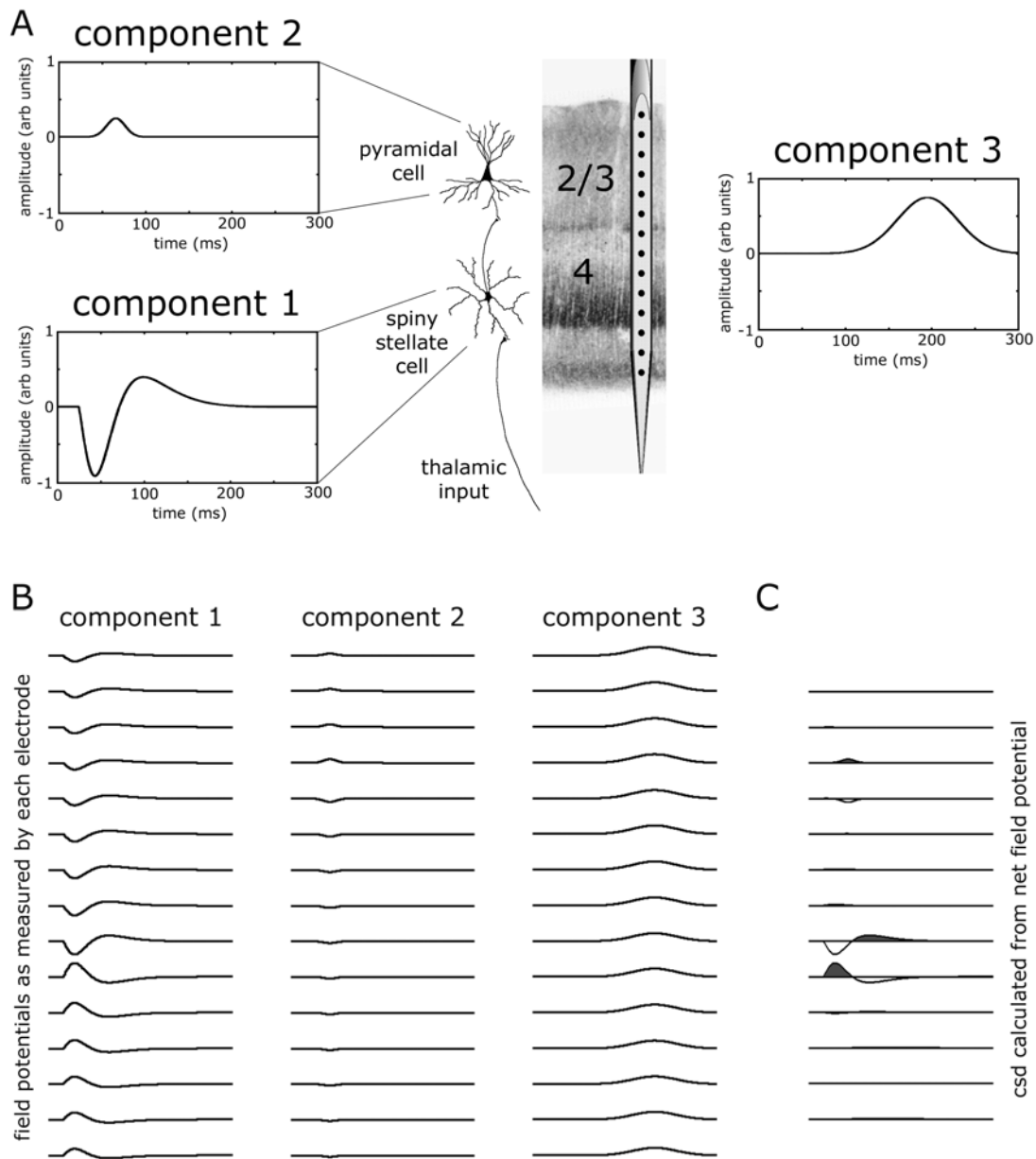


Figure 2 Synthetic data used in the simulations.

The synthetic data utilized in this paper are derived from a model of expected responses in macaque V1 when stimulated by a red-light flash. (A) These data represent electric field potentials recorded from a 15 channel linear-array multielectrode spanning the cortical laminae in V1. The thalamic input activates the spiny stellate cells in layer IV, generating a biphasic field potential (component 1). The feedforward connections from the stellate cells activate the pyramidal cells in the supragranular layers (component 2). Component 3 models a signal generated by a far-field source which volume conducts to the multielectrode array. (B) The noise-free synthetic field potentials generated by these three components are recorded differently by each electrode in the multielectrode array.

Notice that the polarity and amplitude of the recordings depends on the physical positions of the current sources and sinks in the cortical laminae (see C). (C) The current source density (CSD) of the summed field potentials in B is computed using an approximation of their second spatial derivative. The CSD focuses the activity at the location of the current sinks (negative) and sources (positive) relative to the detector positions. This technique clarifies the laminae in which the field potentials are generated. Notice that far-field sources do not appear in the CSD.

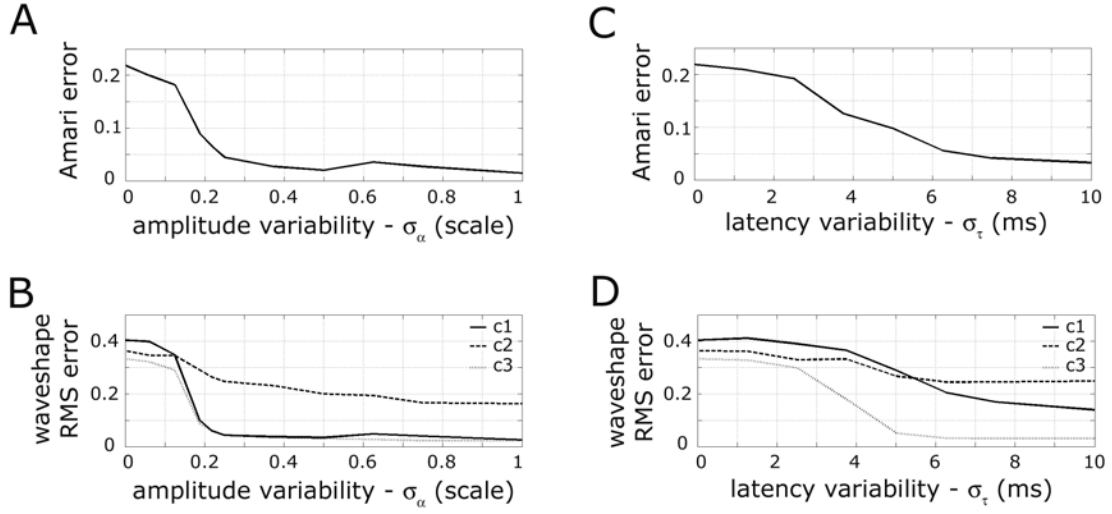


Figure 3 Amplitude and latency variability aids component identification.

(A) The Amari error, which measures the degree of signal separation, decreases with increasing amplitude variability with $\sigma_{amp} \geq 0.25$ being sufficient for signal separation.

(B) The waveshape RMS errors also indicate the importance of amplitude variability.

(C) Amari error decreases with increasing latency variability, with $\sigma_{lat} \geq 7.5$ ms being necessary to achieve effective separation. (D) Increasing latency variability also improves the estimate of the component waveshapes, but not as dramatically as amplitude variability.

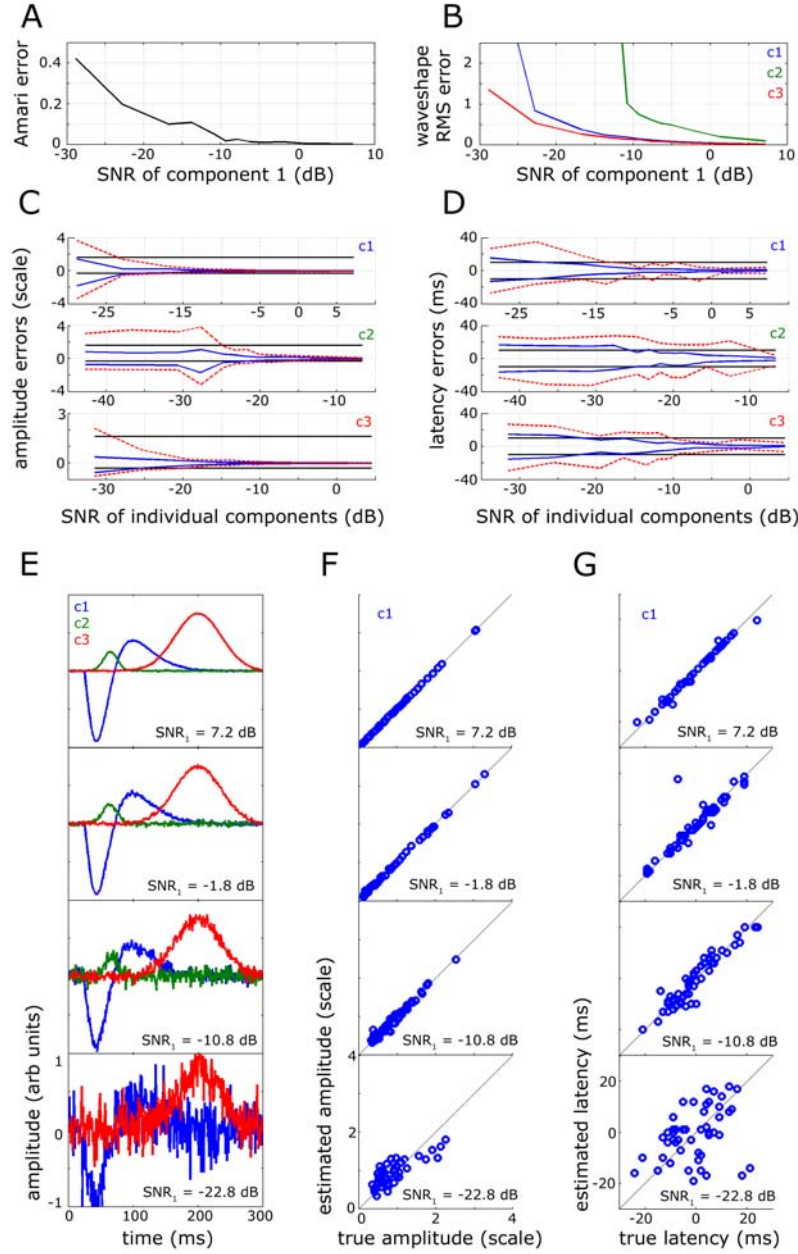


Figure 4 The dVCA algorithm is robust to noise.

(A) The Amari error increases with decreasing SNR (see text). (B) The quality of the wavelshape estimates degrades with decreasing SNR. Note that the graph is drawn with respect to c1 SNR. The wavelshape RMS error reach 1.0, signifying that the deviations in the estimates are on the order of the wavelshape itself, at about -23 dB to -25 dB for the localized components and about -28 dB for the far-field component. (C) The single-trial amplitude estimates are also robustly recovered. The horizontal black lines denote the level of variability in the single-trial quantities. The blue solid and red dashed curves indicate one and two standard deviations, respectively, of the single-trial amplitude estimate errors. When the blue curves are within the black lines, dVCA is performing

better than the standard practice of averaging (see text). High quality estimates are achieved down to -18 dB to -23 dB, and adequate estimates down to -25 dB. (D) Single-trial latency estimates are less well estimated than amplitudes. High-quality estimates are possible down to -8 dB to -12 dB, and adequate estimates down to about -20 dB. (E) These plots demonstrate the quality of the waveshape estimates for four SNR levels. At a c1 SNR of -22.5 dB, the smallest component, c2, was unable to be extracted from the data. (F) Scatter plots of the true c1 single-trial amplitudes versus the estimated c1 single-trial amplitudes demonstrate the quality of amplitude estimates, and show that useful information can be obtained down to an SNR of -22.5 dB. (G) Scatter plots of the true c1 single-trial latencies versus the estimated c1 single-trial latencies demonstrate that latencies can be estimated down to SNR levels below -10.8 dB, but become inaccurate by -22.5 dB.

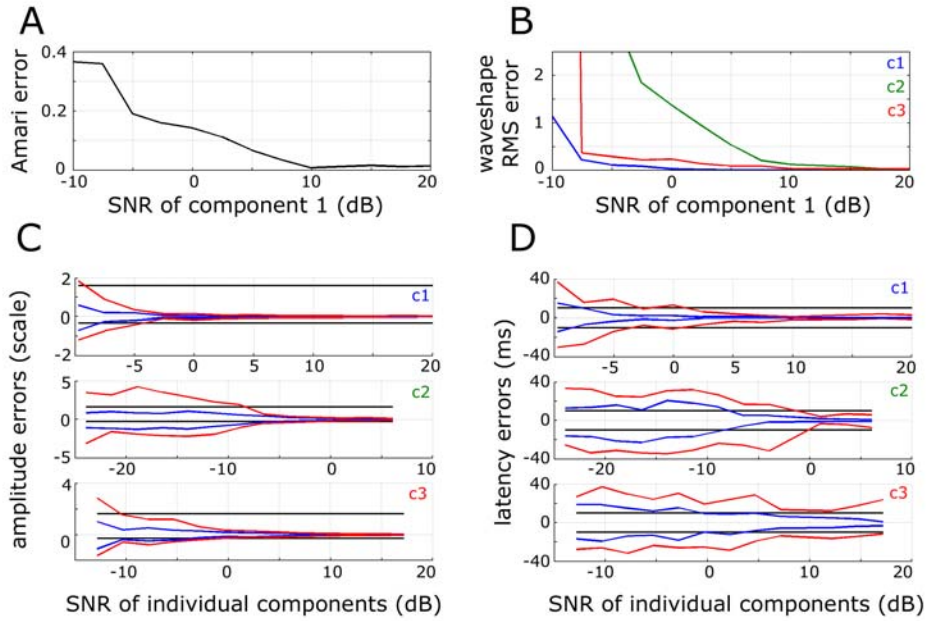


Figure 5 The dVCA algorithm is also robust to correlated far-field noise, although to a lesser degree than independent Gaussian noise.

(A) The Amari error increases with decreasing SNR and reaches the same level of error as in the Gaussian case with 20 dB less noise. (B) The component waveshapes blow up around -10 dB, with the far-field component, c3, being more dramatically affected. (C) The amplitude errors reach unacceptable levels around -5 dB to -8 dB. (D) The latency estimates become unacceptable between -7 and -8 dB for the local sources, whereas the far-field component is even poorly estimated at levels above 5 dB.

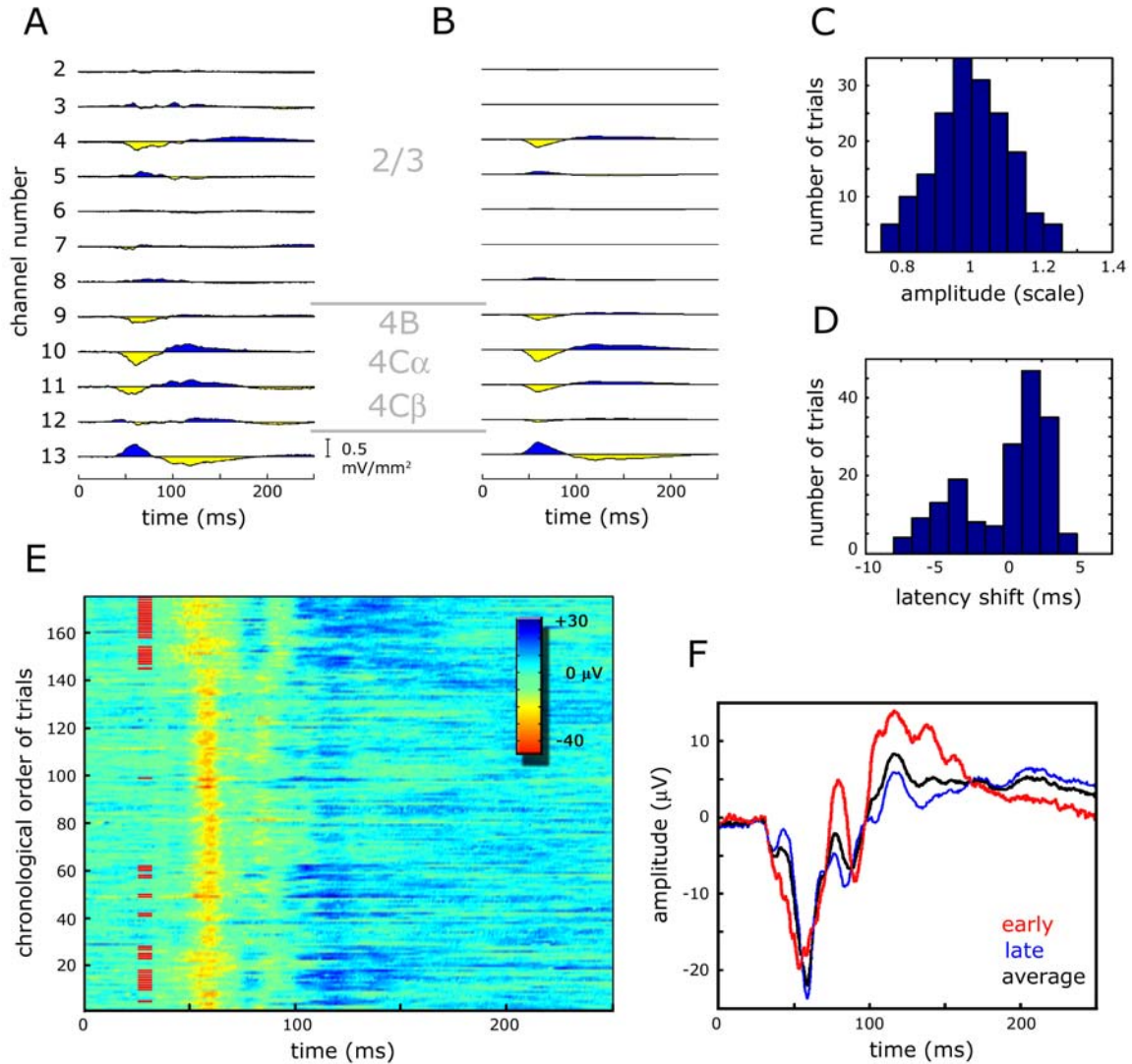


Figure 6 Single component analysis of V1 responses.

(A) CSD of the trial-average ERP shows granular and supragranular activation in macaque V1 in response to a red light flash (number of trials = 171). (B) Estimating a single component with dVCA results in a waveshape with a CSD profile that captures the most prominent responses in the data. (C) A histogram of the single-trial amplitudes of this component shows that the amplitude rarely varies more than $\pm 20\%$. (D) The histogram of the single-trial latencies reveals that there are two response modes: an early response that happens 1/3 of the trials and a late response that happens about 2/3 of the trials. The peak latency difference between these two modes of activation is 6.75 ms. (E) To verify the existence of these two activation modes, this figure shows all 176 trials of the field potential recorded in channel 10. Each trial is represented as a horizontal line with its time-varying color representing the time-varying amplitude of the field potential. Trials designated as belonging to the "early" subset are indicated by the red dashes on the left side of the plot. Note that the "early" trials are characterized by a negative (yellow) field potential onset occurring before the onset seen in the "late" trials. (F) Further verification of this finding is provided by comparing the average FP recorded in channel 10 with the average obtained from the "early" subset and the average obtained from the

“late” subset. While both the early and late responses show initial activation occurring at the same time, the early response’s activation continues to grow, while the late response’s activation decreases before growing again. The difference in latency between the minima of the two sub-averages is 5.5 ms.

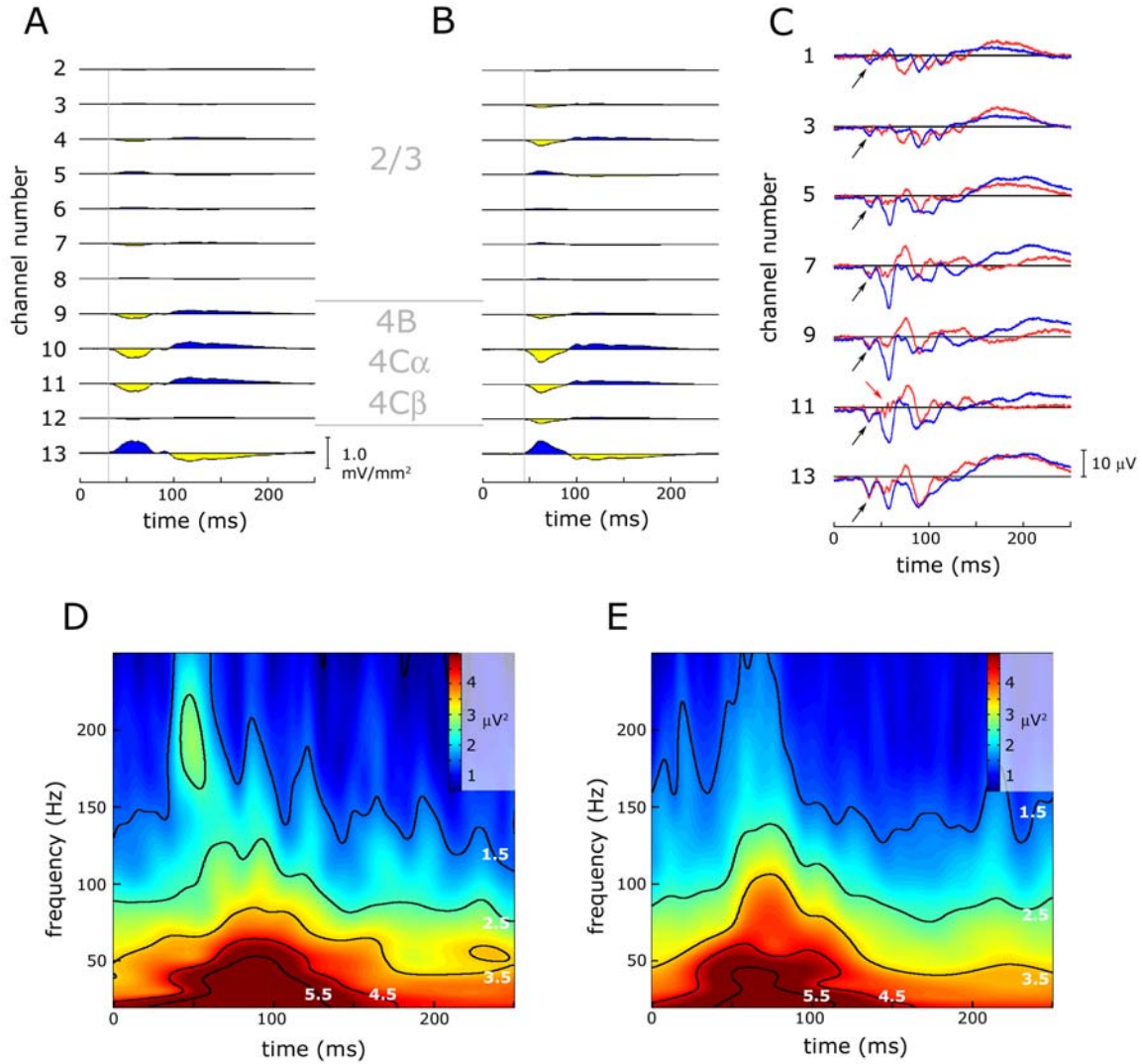


Figure 7 Examination of the early and late responses.

The dataset has been split into two subsets: the early subset and the late subset, and a single component has been re-estimated for each subset. (A) The CSD profile of the early component with a drop line showing onset of the major response at 28 ms. (B) The CSD profile of the late component. The drop line shows its major response onset at 42.5 ms. The waveshapes of the two responses are noticeably different, as are their laminar profiles. (C) The average residual field potentials, computed by subtracting the single-trial model from the single-trial data and averaging over all trials, further demonstrates the differences between these two modes of activation (early-red, late-blue). Only the odd channels are shown. These residuals represent other responses not modeled by the single component in each subset. The black arrows indicate thalamic input, which is time-locked to the stimulus as it is visible in the average. Subsequent activation of the two response types is very different. The late oscillations in channel one after 100 ms are 180 degrees out of phase. The red arrow in channel 11 at about 50 ms shows time-locked ultra-high frequency (UHF) oscillations in the early responses that are not present in the late subset. (D) Wavelet analysis was performed on the residuals to characterize these oscillations and verify that the categorization indicated by the dVCA results is justified.

The residuals in the early subset show a burst of UHF oscillation between 42-57 ms with frequency ranging from 160-220 Hz. (E) These oscillations are completely absent in the late subset.

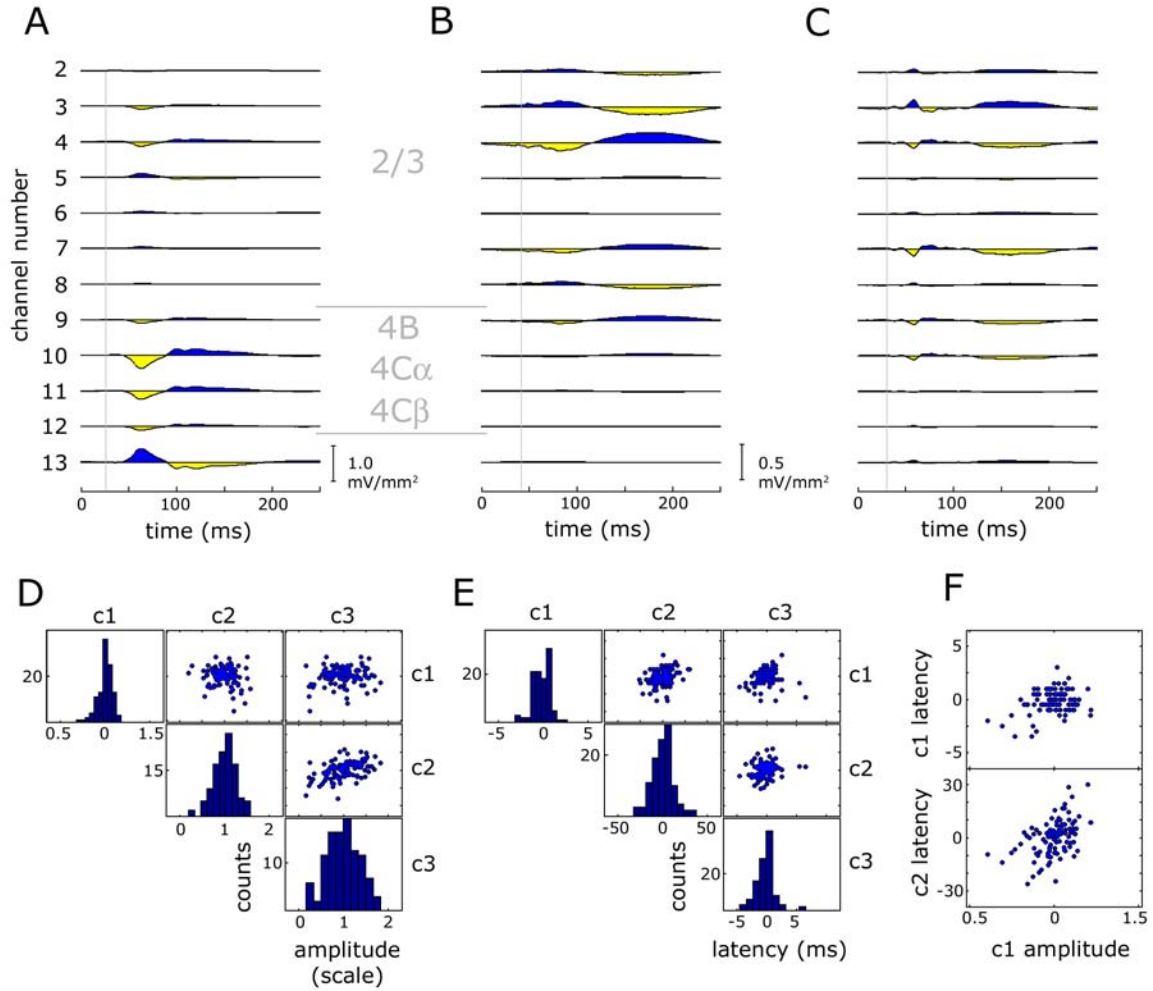


Figure 8 Three components were estimated from the late subset.

(A) The CSD profile of component 1 shows that it represents the granular response with some activation in the supragranular layers. The drop line marks the onset of the response triggered by the thalamic input at 26 ms. (B) Component 2 represents slow-wave activity in the supragranular layers. Its onset is considerably later at 37 ms. (C) Component 3 also describes supragranular activation with a pulse-like activation followed by some slow wave activity. The similarity in the laminar profiles of c2 and c3 suggests that these responses are probably taking place in the same population of cells. Note however that these laminar profiles are not identical as c3 displays some granular activation. (D) The single-trial amplitude histograms as well as amplitude scatter plots are shown. C1 shows very little amplitude variability, whereas that of c2 and c3, while comparable, are not equal. The scatter plot of c2 and c3 amplitudes are correlated with $r = 0.495$ ($p < 10^{-7}$). (E) The single-trial latency histograms and scatter plots are shown. Note that the slow-wave activity in c2 shows great latency variability. The scatter plot of c1 and c2 latencies shows some correlation at $r = 0.327$ ($p < 0.001$). (F) These scatter

plots show the relationship between c1 amplitudes and c1 and c2 latencies with $r = 0.297$ ($p < 0.002$) and $r = 0.483$ ($p < 10^{-7}$), respectively.